# 20.18 Optimization Problems in Air Pollution Modeling

Ivan Dimov, and Zahari Zlatev

**ABSTRACT** Environmental problems are becoming more and more important for modern society, and their importance will certainly be increased in the next century. High pollution levels (high concentrations and/or depositions of certain chemical species) may cause damage to plants, animals, and humans. Moreover, some ecosystems can also be damaged (or even destroyed) when pollution levels are very high. This is why pollution levels must be carefully studied in order to be able to predict the appearance of high pollution levels and/or to decide what can be done to prevent exceeding prescribed critical levels. Mathematical models can successfully be used to resolve these problems. Very often optimization problems have to be applied. The appearance of optimization problems in the field of air pollution modeling and their importance will be discussed in this paper. Some applications of adjoint equations in the treatment of optimization problems arising in air pollution modeling will be considered. We shall present a review of some approaches that are based on the adjoint equations formulation.

## 1. Need for Optimal Environmental Solutions

The control of the pollution levels in different highly developed and densely populated regions of Europe and North America is an important task for modern society. Its importance has been steadily increasing during the last two decades. The need to establish reliable control strategies for air pollution levels will become even more important in the next century. Large scale air pollution models can successfully be used to design reliable control strategies.

The nature of the environmental problems is such that it is highly desirable (and sometimes necessary) to find optimal solutions. This fact can be illustrated by the following simple example. It is very easy to suggest a reduction of all emissions by some large amount (say, by 50 percent everywhere in Europe). Most of the environmental problems would certainly be resolved if such a decision was carried out in practice. However, there is a serious danger of creating some new and even greater problems, because the reduction of the emissions is an expensive process and could cause a great economic crisis in the society if unjustified large reductions were carelessly enforced. This is why optimal solutions are needed in this situation. One has (i) to justify the need for reductions and (ii) to find out where to reduce the emissions and by how much to reduce them. These tasks can be solved successfully by developing and using reliable mathematical models for studying different pollution phenomena (Zlatev, 1995; McRae *et al.*, 1984). These models must satisfy several important requirements (Zlatev *et al.*, 1996):

1. The mathematical models must be defined on large space domains, because the long range transport of air pollution is an important environmental phenomenon and high pollution levels are not limited to the areas where the high emission sources are located.
2. All relevant physical and chemical processes must be adequately described in the models used.
3. Enormous files of input data (both meteorological data and emission data) are needed.
4. The output files are also very big, and fast visualization tools must be used in order to represent the trends and tendencies, hidden behind many megabytes (or even many gigabytes) of digital information, so that even nonspecialists can easily understand them.

## 2. Mathematical Formulation of an Air Pollution Model

All important physical and chemical processes must be taken into account when an air pollution model is to be developed. Systems of partial differential equations (PDEs) are often used to mathematically describe an air pollution model. Consider a three-dimensional space domain

$\Omega$ and assume that $c(r,t) \equiv (c_1(r,t), c_2(r,t), \ldots c_q(r,t))^T$, where $r \equiv (x,y,z) \in \Omega$. Then the PDE systems are of the following type:

$$\frac{dc(r,t)}{dt} - A(r,t)c(r,t) = f(r,t), \tag{1}$$

where

$$A(r,t) \equiv (A_1(r,t), A_2(r,t), \ldots, A_q(r,t))^T, \tag{2}$$

$$A_s(r,t)c_s = -\frac{\partial(uc_s)}{\partial x} - \frac{\partial(vc_s)}{\partial y} - \frac{\partial(wc_s)}{\partial z} + \frac{\partial}{\partial x}\left(K_x \frac{\partial c_s}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial c_s}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_z \frac{\partial c_s}{\partial z}\right) - (\kappa_{1s} + \kappa_{2s})c_s, \tag{3}$$

$$f \equiv (f_1, \ldots, f_q)^T, \tag{4}$$

and

$$f_s = E_s + Q_s(c_1, c_2, \ldots, c_q), \quad s = 1, 2, \ldots, q. \tag{5}$$

The different quantities that are involved in the mathematical model have the following meaning:

- the concentrations are denoted by $c_s$;
- $u$, $v$, and $w$ are wind velocities;
- $K_x$, $K_y$, and $K_z$ are diffusion coefficients;
- the emission sources in the space domain are described by the functions $E_s$;
- $\kappa_{1s}$ and $\kappa_{2s}$ are deposition coefficients;
- the chemical reactions are described by the nonlinear functions $Q_s(c_1, c_2, \ldots, c_q)$.

The nonlinear functions $Q_s$, representing the chemical reactions in which the $s$th pollutant is involved, are of the form:

$$Q_s(c_1, c_2, \ldots, c_q) = -\sum_{i=1}^{q} \alpha_{si} c_i + \sum_{i=1}^{q}\sum_{j=1}^{q} \beta_{sij} c_i c_j, \qquad s = 1, 2, \ldots, q.$$

This is a special kind of nonlinearity (it is seen that the chemical terms are described by quadratic functions), but it is not clear if this property can efficiently be exploited. To the authors' knowledge, it is not exploited in the existing large-scale air pollution models.

The models defined by (1)–(5) are traditionally used to calculate some concentration fields by using both meteorological and emission data as input (Zlatev *et al.*, 1996). This gives an answer to the question: what are the concentration levels and/or the deposition levels caused by the existing emissions under the particular meteorological conditions that take place in the time-period under consideration? However, it is much more important to study the question: how can the concentrations be kept under certain critical levels? Different types of optimization problems must be solved when this and similar questions are to be answered. The formulation of some of the optimization problems arising in air pollution modeling will be studied in the rest of this paper.

## 3. Data Assimilation in Air Pollution Models

Initial concentration fields are needed when the model defined by (1)–(5) is to be treated numerically (and this is, of course, true for any other air pollution model). The problem of finding good initial concentration fields is very difficult. Normally, one starts with some background concentrations and runs the model over a certain sufficiently large period (say, five days) to generate initial concentration fields. It is expected that the initial fields found in this way are good, but there is no guarantee that this is so. Moreover, this is an unnatural solution in the case where an air pollution forecast for the next two–three days is to be produced

(because the time to start-up the model, five days, is much larger than the period of two–three days that is needed for the forecast).

An alternative approach is based on the use of time series measurements to analyze the initial data that is to be used in the air pollution model under consideration. The experiments indicate the remarkable fact that time series of only a few key chemical species are sufficient to produce initial fields (also for species that are not measured; by exploiting their coupling with the measured species).

It must be emphasized here that data assimilation is a very powerful approach. Not only can it be used to analyze the initial data (as sketched above), but also to obtain optimal estimations of (i) emission fields, (ii) deposition rates, and (iii) other model parameters. The use of data assimilation can be considered as applying a sequence of corrections performed at successive times in the treatment of the model. Each correction combines the appropriate background field of results produced by the model with new observations.

Data assimilation has been used successfully in meteorology. The basic ideas can be explained by sketching the application of data assimilation in this field. Before the development of data assimilation techniques in meteorology, the most successful device used was optimal analysis or optimal interpolation. Such a technique is presented in the works of McPherson (1979), Lorenc (1981), and Hollingsworth *et al.* (1985). The advantage of optimal analysis is that it provides a simple and internally consistent procedure for treating a large number of observations with different distributions, nature, and accuracy. However, optimal analysis, as it has been used, only takes into account the dynamics of meteorological processes very indirectly. It is because the temporal evolution of the statistical covariances of the predicted error is represented by a very simple low. This low is taken independently or almost independently of the current state of the wind field which determines the real evolution of the predicted error. In particular, it ignores all effects due to advection. In the paper by Talagrand and Courtier (1987) a variational data assimilation technique is considered. This technique uses the adjoint equations of the model. This approach is based on the Lions' theory of optimal control (see Lions, 1971). Generally speaking, the theory of optimal control deals with the problem of how the output parameters of a given numerical model can be controlled by acting on the input parameters of the model. Among different approaches to control theory the approach of the adjoint equations is one of the most efficient. For meteorological problems this idea was first suggested by Marchuk (1974) to compute the local gradient of a multi-variable function. The computed gradient is then used for performing a descent step in the space of initial conditions and the process is continued until some satisfactory approximation of the initial conditions minimizing the distance function in the corresponding functional space has been fulfilled. This specific application of adjoint equations was implemented in a model by Penenko and Obraztsov (1976), but the feasibility and usefulness of this approach was demonstrated in the works of Derber (1985) and Lewis and Derber (1985). In the works of Le Dimet and Talagrand (1986) and Agoshkov (1994) one can find a rather general presentation of the use of the adjoint equation technique in data assimilation analysis.

There are several different ways of implementing the data assimilation approach in the field of air pollution modeling. Since the mid-eighties an approach based on the principle of optimal control theory has become more and more popular. In this approach the initial values are viewed as control parameters. Then a distance function is defined. This function provides weighted and accumulated distances between the available measurements and the values of the corresponding state variables calculated by the model during a predefined assimilation window. An optimization procedure is applied to minimize the distance function. This requires some knowledge of local gradients with respect to the initial state. To obtain these gradients one has to use the adjoint equation of (1). The procedure is normally called a four-dimensional variational data assimilation.

### *3.1. Introduction of the Distance Function*

Consider a state variable $x$ which is an element of a Hilbert space $\mathcal{H}$. The evolution in time of the state variable $x$ can be expressed by the differential equation:

$$\frac{dx}{dt} = M(x), \tag{6}$$

where $M$ is a nonlinear operator. Denote the measurements by $\hat{x}$ and let

$$\mathcal{O}(x(t)) \stackrel{\text{def}}{=} \tfrac{1}{2}\hat{x}(t) - x(t))O^{-1}(\hat{x}(t) - x(t)),$$

where $O$ is the covariance matrix of the observation and representativeness errors. By using this notation and Lagrangian multipliers $\lambda(t)$, the distance function can be defined by

$$\mathcal{J}(x(t)) \stackrel{\text{def}}{=} \int_{t_0}^{t_N} \left[ \mathcal{O}(x(t)) + \left\langle \lambda(t), \frac{dx(t)}{dt} - Mx(t) \right\rangle \right] dt, \tag{7}$$

where the inner product $\int_\Omega f(x)g(x)dx$ of two functions $f$ and $g$ in the Hilbert space $\mathcal{H}$ is denoted by $\langle f, g \rangle$. It should be mentioned here that the background concentrations can also be used in the definition of the distance function by adding an extra term

$$\tfrac{1}{2}(x_b - x(t_0))^T B^{-1}(x_b - x(t_0))$$

to the right-hand-side of (7).

### *3.2. Calculating the Gradient of the Distance Function*

A perturbation equation (or a variational equation), corresponding to (6), can be defined by

$$\frac{d(\delta x)}{dt} = M'(\delta x), \tag{8}$$

where $\delta x$ is some small deviation of $x$ and $M'$ is the tangential linear operator of $M$. Denote by $g_1$ and $g_2$ the inner products:

$$g_1 = \left\langle \delta\lambda(t), \frac{dx(t)}{dt} - Mx(t) \right\rangle,$$

$$g_2 = \left\langle \frac{\delta\lambda(t)}{dt}, \frac{d(\delta x(t))}{dt} - M'\delta x(t) \right\rangle.$$

In this notation, the gradient (the variation) of the distance function introduced in the previous paragraph can be defined as follows:

$$\delta\mathcal{J} \stackrel{\text{def}}{=} \int_{t_0}^{t_N} [\langle \delta\mathcal{O}, \delta x(t)) \rangle + g_1 + g_2] dt.$$

After some modifications of this equation (the most important of which is the traditionally used integration by parts), the application of the extremal principle $\delta\mathcal{J} = 0$ results in the adjoint equation:

$$-\frac{d\lambda(t)}{dt} - M'^*\lambda(t) = O^{-1}[\hat{x}(t) - x(t)],$$

where the adjoint operator $M'^*$ is defined by $\langle y, M'z \rangle = \langle M'^* y, z \rangle$.

In order to estimate the evolution in time of the given initial perturbation of the data at time $t_0$ we have to integrate equation (8):

$$\delta x(t_n) = R(t_n, t_0)\delta x(t_0), \quad n = 1, ..., N, \tag{9}$$

where $R(t_n, t_0)$ is the resolvent operator of $M'$. In general, we have to use some numerical integration scheme in order to calculate the resolvent operator $R(t_n, t_0)$. If this is done, then the resolvent operator is expressed as

$$R(t_n, t_0) = \prod_{i=0}^{n-1} R(t_{i+1}, t_i), \quad n = 1, ..., N,$$

where $R(t_{i+1}, t_i)$ is a sufficiently accurate numerical operator for stepwise integration.

Let us now consider the operator $S$ as the resolvent of $M'^*$. From the duality principle it follows that

$$\langle \lambda(t_{t_N}), R(t_N, t_0)\delta x(t_0)\rangle = \langle S(t_0, t_N)\lambda(t_N), \delta x(t_0)\rangle$$

and

$$\nabla_{x(t_0)}\mathscr{J} = \sum_{m=1}^{N}\prod_{i=0}^{m-1} S(t_i, t_{i+1})\nabla_x \mathscr{O}(t).$$

It remains to be shown that $\nabla_{x(t_0)}\mathscr{J} = \lambda(0)$. After defining the backward initial condition $\lambda(t_N) = 0$ and applying the duality principle, one can get

$$-\frac{\partial}{\partial t}S(t, t')\nabla_{x(t')}\mathscr{O}(t') = -\frac{d\lambda(t)}{dt} = -M'^*S(t, t')\nabla_{x(t')}\mathscr{O}(t'). \tag{10}$$

After the integration of (10) one can get

$$\lambda(t) = \int_t^{t_N} S(t, t')\nabla_{x(t')}\mathscr{O}(t')dt'.$$

The computations concerning the calculation of the values of the distance function $\mathscr{J}(x(t_0))$ and of its gradient $\delta\mathscr{J}(x(t_0))$ can be carried out in two steps.

- *Forward step.* Some approximations of the state variable $x$ and $\delta x$ are calculated, from (6) and (8), respectively, for $t = t_1, t_2, ..., t_N$. The values of $x$ are used to calculate the values of the distance function $\mathscr{J}(x(t_0))$ in the same points.
- *Backward step.* The adjoint equation (or some transformed form of this equation) is integrated backward for $t = t_{N-1}, t_{N-2}, ..., t_0$ to obtain values of the perturbed state variable $\lambda(t)$ starting with $\lambda(t_N) = 0$. The values of $\lambda(t)$ and $\delta x(t)$ are used to calculate values of $\delta\mathscr{J}$.

The values obtained for $\mathscr{J}$ and $\delta\mathscr{J}$ are needed in the minimization procedure which will be outlined in the next section.

### 3.3. *Minimizing the Distance Function and Beneficial Impact from the Assimilation of the Selected Species*

Different standard optimization algorithms can be applied to minimize the distance function. The algorithm proposed by Broyden–Fletcher–Goldfarb–Shanno (BFGS) is often used in this field in its limited memory version. It should be stressed, however, that nonnegative solutions are required. Therefore, it is necessary to insert some constraints which ensure nonnegative solutions. A procedure of this kind is proposed by Bertsekas (1982).

Let us describe briefly a practical minimization procedure of the distance function in order to demonstrate to what extent measurements of a limited set of species can reduce the manifold of possible initial perturbations to fit to the observations. Let the observations of species $1, ..., p$ out of $p \leqslant q$ be available at $N$ time-steps. Denote by $\delta_p\hat{\mathbf{x}}(t_n)$ the following vector:

$$\delta_p\hat{\mathbf{x}}(t_n) = (\delta\hat{\mathbf{x}}_1(t_n), ..., \delta\hat{\mathbf{x}}_p(t_n))^T, \quad n = 1, ..., N, \tag{11}$$

where $p$ is the number of the components of the vector and $\delta\hat{\mathbf{x}}_i(t_n) = \hat{\mathbf{x}}_i(t_n) - x_i(t_n)$, $i = 1, 2, ..., p$. The system (9) can be rewritten as

$$\delta_p\hat{\mathbf{x}}(t_n) = [\mathbf{I}, \mathbf{0}]R(t_n, t_0), \quad n = 1, ..., N, \tag{12}$$

where $[\mathbf{I}, \mathbf{0}] \in \mathbb{R}^{p\times q}$ is the truncation operator composed of the identity matrix $\mathbf{I} \in \mathbb{R}^{p\times p}$ and the zero matrix $\mathbf{0} \in \mathbb{R}^{p\times(q-p)}$ (for the case $q > p$). In system (12) the rows of unobserved species are truncated. It is convenient to comprise the matrices of the resolvent operator and the truncation operator in a matrix $\mathscr{R}(t_n, t_0)$

$$\delta_{Np}\hat{\mathbf{x}}(t) = \mathscr{R}(t_n, t_0)\delta_{Np}x(t_0), \quad n = 1, ..., N, \tag{13}$$

where $\mathscr{R}(t_n, t_0) \in \mathbb{R}^{Np\times q}$, $\delta_{Np}x(t_0) \in \mathbb{R}^{Np}$, and

$$\delta_{Np}\hat{\mathbf{x}}(t) = (\{\delta_p\hat{\mathbf{x}}(t_1)\}^T, ..., \{\delta_p\hat{\mathbf{x}}(t_N)\}^T)^T \in \mathbb{R}^{Np}$$

(the row vectors $\delta_p\hat{\mathbf{x}}(t_n)$, $n = 1, ..., N$ are defined by (11)).

It is easy to see that for $Np < q$ the system (13) is underdetermined. In the case where $Np > q$ the system (13) is overdetermined. In this case one can estimate the condition numbers of the matrix $\mathscr{R}'(t_n, t_0) \in \mathbb{R}^{q \times q}$, part of $\mathscr{R}(t_n, t_0)$, which is a square matrix of size $q \times q$, that is $cond(\mathscr{R}'(t_n, t_0))$ with any value of $mh_i = q$, where $h_i$ is any combination of measured species and $m$ is the number of the time-steps. The condition numbers $cond(\mathscr{R}'(t_n, t_0))$ for different combinations give us an estimate of what beneficial impact may be expected from the assimilation of the selected species $i$ and observation times $m$. These condition numbers can provide some evidence of reasonable weighting of the observations, which is an important task. For more details see Elbern *et al.* (1997).

## 4. Keeping the Concentrations in a Given Sensitive Area Under a Prescribed Level

The problem of environmental protection of a given sensitive area is an optimization problem. This problem can be formulated as follows. Let a sensitive point $r_1 \in \Omega$ be given. Assume that the concentrations at point $r_1 \in \Omega$ in time $t_1 \in [0, T]$ should not exceed some critical (for point $r_1$) level $C$. This means that the requirement

$$c(r_1, t_1) \leqslant C \tag{14}$$

must be satisfied. Assume also that some source of a prescribed power $E(t)$ is to be located somewhere in $\Omega$. The problem is to find a subdomain $\omega$ of $\Omega$ where the source is located, so that condition (14) is satisfied. This problem could also be considered as a problem of optimal planning in an effort to achieve sustainable development.

It is convenient to simplify the problem in order to facilitate the understanding of the main principles which can be used to find the fundamental solution of the adjoint problem. This can be done as follows. Consider the solution of $1D-$transport diffusion problem:

$$\left(\frac{\partial}{\partial t} - A\right)c = \frac{\partial c}{\partial t} - \frac{\partial}{\partial x}(\alpha x c) - D\frac{\partial^2 c}{\partial x^2} = E(t)\delta(x - x_0), \tag{15}$$

$$c = c_0(x), \quad \text{for} \quad t = 0. \tag{16}$$

Assume that $c(x, t) \in W_1^2$, where $W_1^2$ is a Sobolev space, $c(x, t) > 0$, $D > 0$, $\alpha = \text{const}$, $-\infty < x < \infty$, $t \in [0, T]$, and the conditions providing the existence of the unique solution of the problem (15)–(16) in a weak formulation are fulfilled.

It is not possible to directly obtain an analytical solution to the problem defined by (15)–(16), due to presence of the term $E(t)\delta(x - x_0)$. In addition, standard numerical methods used to solve the equation (15) are not efficient because there is a $\delta$-function in (15).

The condition $c(x, t) \in W_1^2$ implies that $\int_0^T dt \int_{-\infty}^{\infty} c^{(2)}(x, t)dx < \infty$, which is a natural condition.

In fact, the operator used in (15) contains a rather artificial wind field defined by $u = \alpha x$. Nevertheless, we demonstrate in this way how the approach under consideration can be used for problems with a nonconstant advection.

The problem of finding the subdomain $\omega$ is not so complicated if the solution of the adjoint formulation of the original problem (15)–(16) is used. The conditions formulated here will be relaxed in Section on Application to Other Problems.

### *4.1. Adjoint formulation*

Multiply equation (15) by a function $c^*$, which will be defined later (see the paragraph after equation (25–26)). Integrate the resulting equation in time and space:

$$\int_0^T dt \int_{-\infty}^{\infty} c^* \left(\frac{\partial c}{\partial t} - \frac{\partial}{\partial x}(\alpha x c) - D\frac{\partial^2 c}{\partial x^2}\right)dx = \int_0^T E(t)dt \int_{-\infty}^{\infty} c^* \delta(x - x_0)dx. \tag{17}$$

Assume that $c = c^* = 0$ when $x \to \infty$ or $x \to -\infty$. Apply integration part by part to (17). The following results will be obtained in this way:

$$\int_0^T dt \int_{-\infty}^{\infty} c^* \frac{\partial c}{\partial t} dx = \int_{-\infty}^{\infty} cc^* dx|_{t=0}^{t=T} - \int_0^T dt \int_{-\infty}^{\infty} c \frac{\partial c^*}{\partial t} dx$$

$$= -\int_{-\infty}^{\infty} c_0(x) c_0^*(x) dx - \int_0^T dt \int_{-\infty}^{\infty} c \frac{\partial c^*}{\partial t} dx, \tag{18}$$

$$\int_0^T dt \int_{-\infty}^{\infty} c^* \frac{\partial^2 c}{\partial x^2} dx = \int_0^T \left( c^* \frac{\partial c}{\partial x} - c \frac{\partial c^*}{\partial x} \right) dx|_{x=-\infty}^{\infty} + \int_0^T dt \int_{-\infty}^{\infty} c \frac{\partial^2 c^*}{\partial x^2} dx, \tag{19}$$

and

$$\int_0^T dt \int_{-\infty}^{\infty} c^* \frac{\partial (xc)}{\partial x} dx = \int_{-\infty}^{\infty} xcc^* dx|_{t=0}^{t=T} - \int_0^T dt \int_{-\infty}^{\infty} xc \frac{\partial c^*}{\partial x} dx$$

$$= -\int_{-\infty}^{\infty} xc_0(x) c_0^*(x) dx - \int_0^T dt \int_{-\infty}^{\infty} xc \frac{\partial c^*}{\partial x} dx. \tag{20}$$

From (18), (19), and (20), one can obtain:

$$\int_0^T dt \int_{-\infty}^{\infty} c \left( -\frac{\partial c^*}{\partial t} + \alpha x \frac{\partial c^*}{\partial x} - D \frac{\partial^2 c^*}{\partial x^2} \right) dx - \int_{-\infty}^{\infty} (1-x) c_0(x) c_0^*(x) dx = \int_0^T E(t) c^*(x_0, t) dt. \tag{21}$$

Assume that function $c^*$ satisfies the following adjoint equation:

$$-\frac{\partial c^*}{\partial t} + \alpha x \frac{\partial c^*}{\partial x^*} - D \frac{\partial^2 c}{\partial x^2} = \psi(x, t),$$

with an initial condition $c^*(x, T) = 0$, where the function $\psi(x, t)$ will be defined later.

It is possible to rewrite (21) in the following form:

$$\int_0^T dt \int_{-\infty}^{\infty} \psi(x, t) c(x, t) dx = \int_0^T E(t) c^*(x_0, t) dt + \int_{-\infty}^{\infty} (1-x) c_0(x) c_0^*(x) dx = J \tag{22}$$

In fact, functions $c(x, t)$ and $c^*(x, t)$, as well as functional $J$, depend on parameters $x_0$, $x_1$, and $t_1$, so that $c(x, t) = c(x, t, x_0, x_1, t_1)$, $c^*(x, t) = c^*(x, t, x_0, x_1, t_1)$, and $J = J(x_0, x_1, t_1)$. When we do not use these parametric dependences, they will be omited.

It is necessary to find an efficient numerical method to compute $J$. One can choose the function $\psi(x, t)$ in the form of a product of two $\delta$-functions: $\psi(x, t) = \delta(x - x_1)\delta(t - t_1)$.

Let us also assume that $E(t)$ can be represented as

$$E(t) = e(t)H(t), \tag{23}$$

where $H(t)$ is a Heaviside function ($H(t) = 1$ for $t > 0$ and $H(t) = 0$ for $t \leqslant 0$). The second term of the right-hand-side of (22) vanishes when such an assumption is made.

Let us now consider the following linear functional of the solution of the original problem (15)–(16):

$$J = \int_0^T dt \int_{-\infty}^{\infty} \psi(x, t) c(x, t) dx. \tag{24}$$

Under the above assumptions, we have $J = \int_0^T dt \int_{-\infty}^{\infty} c(x, t) \delta(x - x_1) \delta(t - t_1) dx = c(x_1, t_1)$.

Obviously, the value of the functional (24) is equal to the value of the concentration in point $x = x_1$ at moment in time $t = t_1$. From the representation (22) it follows that instead of solving the original problem (15)–(16) one can solve the adjoint problem for $c^*$:

$$-\frac{\partial c^*}{\partial t} + \alpha x \frac{\partial c^*}{\partial x} - D \frac{\partial^2 c^*}{\partial x^2} = \delta(x - x_1)\delta(t - t_1) \tag{25}$$

$$c^* = 0, \quad \text{for} \quad t = T. \tag{26}$$

From the solution of the problem formulated by (25)–(26) one can estimate the functional (22) as a function depending of a parameter $x_0$ and solve the problem under consideration. In fact, if the function $c^*(x, t, x_0, x_1, t_1)$ is the solution of the adjoint problem (25)–(26), then the presentation (22) defines a function of $x_0 - J(x_0, x_1, t_1)$. Now, the set $\omega$ of points $x_0$ can be determined by solving the inequality $J(x_0, x_1, t_1) \leqslant C$. That is the solution of the problem under consideration.

### *4.2. Evaluating the functionals*

One needs an efficient method for evaluating the linear functionals

$$\int_0^{t_i} e(t)c^*(x_0, t)dt \tag{27}$$

of the solution to the adjoint problem (25)–(26) ($e(t)$ is defined by (23)).

It is efficient to apply a statistical numerical method in the evaluation of (27). This is because the complexity of the statistical numerical methods for computing a linear functional of the solution is very low; of the same order as the complexity of the solution at only one point of the domain (see Sobol, 1973; Dimov and Tonev, 1993). The statistical methods give statistical estimates for the functional of the solution by performing samples of a certain random variable, the mathematical expectation of which is the desired functional. These methods have proved to be very efficient in solving multi-dimensional problems in composite domains (Curtiss, 1954, 1956; Hammersley and Handscomb, 1964; Sobol, 1973; Ermakov *et al.*, 1984). Moreover, it has been shown that for some problems (including one-dimensional ones) the statistical methods have a better convergence rate in the corresponding functional spaces than the optimal deterministic methods in the same functional spaces (Bahvalov, 1964; Dimov and Tonev, 1989, 1993b). It is also very important that the statistical methods are very efficient when parallel or vector processors or computers are available, because the above mentioned methods are inherently parallel and have loose dependencies. They are also well vectorizable. By using powerful parallel computers, it is possible to apply a Monte Carlo method or a particle-tracking method for evaluating large-scale nonregular problems which are sometimes difficult to solve by traditional numerical methods.

The difficulties with finite difference and finite element methods are caused by the nonregularity of the right-hand-side of equation (25) and by artificial diffusion (or diffusion caused by the numerical scheme). This explains why the functional (27) can be evaluated by using a statistical numerical method when $A$ is a general linear operator. In our case, the fundamental solution of the adjoint equation can be used.

In fact, introduce a new variable $t^* = T - t$, where $t^* \in [0, T]$ and present the adjoint problem in the following form:

$$\frac{\partial c^*}{\partial t^*} + \alpha x \frac{\partial c^*}{\partial x} - D\frac{\partial^2 c^*}{\partial x^2} = \delta(x - x_1)\delta(T - t^* - t_1) \tag{28}$$

$$c^* = 0, \quad \text{for} \quad t^* = 0, \quad \forall x.$$

The fundamental solution of problem (28) can be found by using the techniques discussed by Bitzadze (1982) and Tikhonov and Samarski (1977):

$$\hat{c}^*(x, t) = \frac{1}{[\pi a(t)]^{1/2}} \exp\left\{ - \frac{[x - b(t)]^2}{a(t)} \right\}, \tag{29}$$

where $a(t) = \frac{2D}{\alpha}(1 - e^{-2\alpha t})$ and $b(t) = x_0 e^{-\alpha t}$.

The solution of the adjoint problem can be expressed in the following form:

$$c^*(x, t^*) = \int_0^T dt^* \int_{-\infty}^{\infty} \hat{c}^*(x - \xi, t^* - \tau)\delta(\xi - x_1)\delta(T - \tau - t_1)d\xi d\tau. \tag{30}$$

The result of integration (30) is the function defined by (29) with new arguments:

$$c^*(x-x_1,t_1-t)=\frac{1}{[\pi a(t_1-t)]^{1/2}}\exp\left\{-\frac{[x-x_1-b(t_1-t)]^2}{a(t_1-t)}\right\}. \tag{31}$$

Substituting (31) into the functional (27) one can get:

$$\int_0^{t_1} e(t)\frac{1}{[\pi a(t_1-t)]^{1/2}}\exp\left\{-\frac{[x_0-x_1-b(t_1-t)]^2}{a(t_1-t)}\right\}dt=J(x_0,x_1,t_1),$$

Under the assumptions formulated in this section the problem has been reduced to a problem of numerical integration. We have to evaluate a 1-D integral in our particular case. For this purpose we do not need to use a special statistical numerical method since the problem is simplified. The proposed numerical method is based on a numerical integration scheme for evaluating the last parameterized functional. Using a Newton–Cotes quadrature one can get

$$J=\sum_{i=0}^{k-1} e_i\frac{1}{[\pi a(t_1-\tau_i)]^{1/2}}\exp\left\{-\frac{[x_0-x_1-b(t_1-\tau_i)]^2}{a(t_1-\tau_i)}\right\}(\tau_{i+1}-\tau_i)+O(\Delta\tau), \tag{32}$$

where $\Delta\tau=\max_i(\tau_{i+1}-\tau_i)$ is the maximal time-step and $e_i=e(\tau_i)$. The numerical scheme (32) is of the order $O(\Delta\tau)$. It is possible to apply more complicated schemes of high order accuracy (say, $O((\Delta\tau)^2)$, or $O((\Delta\tau)^4)$). But from computational point of view, when the integrand is a sufficiently smooth function, which corresponds to the case under consideration, it is better to use an adaptive numerical scheme. In our practical computations an adaptive numerical scheme is used. The integration starts with a regular coarse grid with a low number $k$ of grid points. After evaluating the first coarse approximation to the functional (32), the total and local a posteriori variances are estimated. The information obtained is used for refinement of the mesh—in the subregions in which the local variance is too large the mesh is refined, such that after each step of refinement the number of nodes (mesh-points) increases two times. The refinement procedure is continued until the required accuracy of the integration is reached or until the maximum number of evaluations is executed.

### *4.3. Application to Other Problems*

The approach studied can be easily applied to transport problems with another linear operator $A$. One can for example include deposition terms $kc$ in the operator considered in the previous subsections:

$$\left(\frac{\partial}{\partial t}-A\right)c=\frac{\partial c}{\partial t}-u\frac{\partial c}{\partial x}+\kappa c-D\frac{\partial^2 c}{\partial x^2}=E(t)\delta(x-x_0), \tag{33}$$

$$c=c_0(x),\quad \text{for}\quad t=0, c=0,\quad \text{for}\quad x\to\infty,\quad \text{and}\quad x\to-\infty.$$

Using the technique of the adjoint equations described above one can reconstruct the location and power of a steady state point source from two observations $c(x_1, t_1)$ and $c(x_2, t_2)$. A possible application would be the identification and characterization of a pollutant source in an inaccessible area. This formulation leads to a well conditioned problem and can be generalized to a large class of more complicated cases (see Dimov *et al.*, 1996).

Here we consider the following

**Problem**. *Find the place of location $x_0$ and the power $e$ of the source $E(t)=E=eH(t)$ ($H(t)$ is a Heaviside function) if two measurement data $c(x_1,t_1)$ and $c(x_2,t_2)$ are available (it is assumed, that $x_1\neq x_2$).*

Let the power of source $e$, as well as its location $x_0$ be unknown. Two measurement data points—$c(x_1, t_1)$ and $c(x_2, t_2)$ are assumed to be available. One can then define two functions of $x_0$, $F_1(x_0)$ and $F_2(x_0)$, using the following formal presentations:

$$F_i(x_0)=\frac{\hat{J}(x_0,x_i,t_i)}{c(x_i,t_i)},\quad i=1,2,$$

where

$$\hat{J}(x_0, x_i, t_i) = \frac{J(x_0, x_i, t_i)}{e}, \quad i = 1, 2.$$

(It is possible to do this because $e$ is strongly positive.)

One can see, that $x_0$ is the solution of equation

$$F_1(x) = F_2(x),$$

being the point of the location of the source of pollutants. The value of the power of the source $e$ can be determined using the expression:

$$e = \frac{1}{F_i(x_0)}.$$

The problem consists of using an efficient numerical method for evaluating the functional

$$\int_0^{t_i} ec^*(x_0, t)dt.$$

Such a method for finding the place of location and the power of the sources for problem (33) is formulated and experimented with in Dimov *et al.* (1996, 1997). Computational results are given in Dimov *et al.* (1996) for different values of $D/u$, where $D$ is the diffusion coefficient and $u$ is the wind velocity (see (33)).

## 5. Optimization Problems in the Theory of Measurements and Planning the Experiments

Suppose that $n$ functionals $J_{p_i}$, $i = 1, ..., n$, are available. For every functional $J_{p_i}$ we consider the corresponding weight function $u^*_{p_i}$, $i = 1, 2, ..., n$ for a nonperturbed problem. (It should be emphasized that in this section the functionals $J_{p_i}$ are measurements, $u$ and $v$ are weight functions (not wind velocities), and $p_i$ are also functions.) We consider a model in which the operator $L$, as well as the domain of definition are supposed to be known. Consider the following $n$ problems:

$$L^*u^*_{p_i} = p_i, \quad i = 1, 2, ..., n.$$

Let us find $n$ weight functions $u^*_{p_i}$, $i = 1, 2, ..., n$ and solve the following problem:

$$Lu = E. \tag{34}$$

The nonperturbed operator $L$ is adjoint to the operator $L^*$. In our consideration, we assume that $u \in U$ and $u^* \in U^*$, where $U$ and $U^*$ are the domains of definition of the operators $L$ and $L^*$, respectively.

The goal is to reconstruct the coefficients $\alpha'_k$ and $\beta'_k$ in the expression, by which the perturbed operator $L'$ is defined:

$$L' = \sum_{k=1}^{m} \{\alpha'_k A_k + B_k(\beta'_k C_k)\}, \tag{35}$$

where $A_k$, $B_k$, and $C_k$ are simple linear operators (for example, operators of differentiation, integration, and so on), $\alpha'_k(x)$ and $\beta'_k(x)$ are perturbation coefficients.

Use the theory of small perturbations to construct $n$ formulas:

$$(u^*_{p_i}, \delta Lu) = -\delta J_{p_i}, \quad i = 1, 2, ..., n, \tag{36}$$

where $\delta L$ is the difference between the perturbed operator $L'$ and the nonperturbed operator $L$.

Assume that the operator $L$ can be presented as

$$L = \sum_{k=1}^{m} \{\alpha_k A_k + B_k(\beta_k C_k)\}, \tag{37}$$

where $\alpha_k(x)$ and $\beta_k(x)$ are unknown coefficients which are usually approximately known for nonperturbed (model) problem.

Using expressions (37) and (35) one can get

$$\delta L = \sum_{k=1}^{m} \{\delta\alpha_k A_k + B_k(\delta\beta_k C_k)\}, \tag{38}$$

where

$$\delta\alpha_k = \alpha_k^{'} - \alpha_k \quad \text{and} \quad \delta\beta_k = \beta_k^{'} - \beta_k.$$

Substituting (38) into (36) one can get the following system of equations:

$$\sum_{k=1}^{m} [(u_{p_i}^*, \delta\alpha_k A_k u) + (B_k^* u_{p_i}^*, \delta\beta_k C_k u)] = -\delta J_{p_i}, \tag{39}$$

$i = 1, 2, ..., n.$

Now the problem consists of parameterization of the variances $\delta\alpha_k$ and $\delta\beta_k$. It has been shown (Marchuk, 1980) that if $\delta\beta_k = 0$ and $\delta\alpha_k$ are constants, then (39) is reduced to the linear algebra problem

$$\Lambda y = F, \tag{40}$$

where $\Lambda$ is a matrix with elements $\lambda_{ik} = (u_{p_i}^*, A_k u)$, $F$ is a vector with components $-\delta J_{p_i}$ and $y$ is a vector with components $\delta\alpha_k$.

If the number of functionals (measurement data) $n$ is equal to the number of varied coefficients $\alpha_k$, then the system allows us, in principle, to determine $\delta\alpha_k$. If $n$ is greater than $m$ the solution could be found by minimizing the following quadratic functional:

$$\|\Lambda y - F\|_2^2 = \min, \tag{41}$$

where $\|\cdot\|_2^2$ is the Euclidean norm.

This optimization problem can be efficiently solved using a gradient optimization procedure like the procedure described in Section 3.2.

This approach could be extended for the case when $\delta\alpha_k(x)$ and $\delta\beta_k(x)$ are functions. In this case, the solution of the inverse problem could be found using the parameterization approach. In this approach we suppose that as a result of the statistical and correlation analysis, orthogonal systems of functions $u_{k\cdot l}(x)$ and $v_{k\cdot l}(x)$ could be found. Using such functions one can obtain good approximations to functions $\alpha_k$ and $\beta_k$ for a sufficiently small $n(k)$, ($n$ depends on $k$), so that

$$\delta\alpha_k(x) = \sum_{l=1}^{n(k)} a_{k,l} u_{k,l}(x), \tag{42}$$

$$\delta\beta_k(x) = \sum_{l=1}^{n(k)} b_{k,l} v_{k,l}(x), \tag{43}$$

where $a_{k,l}$ and $b_{k,l}$ are unknown coefficients. Substituting (42) into (39) one can get

$$\sum_{k=1}^{m}\sum_{l=1}^{n(k)} [a_{k,l}[u_{p_i}^*, u_{k,l} A_k u) + b_{k,l}(B_k^* u_{p_i}^*, v_{k,l} C_k u)] = -\delta J_{p_i} \tag{44}$$

$i = 1, 2, ..., n.$

It is easy to see that system (44) is equivalent to system (40). Then one could again consider the minimization problem (41).

Consider the problem of planning a complicated experiment in environmental science. Let us formulate the following

**Problem**. *Consider a set of all possible* (*practically realizable*) *experiments and choose the subset of it which is the best, from the point of view of solving the given inverse problem of reconstruction of the parameters of the environment* (*that is coefficients of the equations*).

Such a problem is very difficult and does not have a solution in the general case. But there are some special approaches to this problem. Suppose that one considers the model of nonperturbed problem (34) before doing the experiment and that one describes linear functionals of the solution. One can find the necessary accuracy for measuring the functionals by taking into account the information about measurement accuracy. Assume that some prescribed requirements to the accuracy of the measurements are fulfilled, i.e. $\|\delta J_{p_i}\| < \varepsilon$, where $\varepsilon > 0$ is a given small value. Then one could consider different sets of experiments and choose the set which leads to the matrix $\Lambda$ defined by (40) with the best condition number (i.e. the $\Lambda$ with the smallest $\|\Lambda\|_2 \|\Lambda^\dagger\|_2$, where $\|\Lambda\|_2$ is any Euclidean norm of the matrices $\Lambda$ and $\Lambda^\dagger$ is the pseudo-inverse of the matrix $\Lambda$ defined by $\Lambda^\dagger = (\Lambda^T\Lambda)^{-1}\Lambda^T$). In fact, by all this we are assuming that the system of linear algebraic equations can be solved and the solution found is sufficiently accurate. Such a plan for the experiments (the subset of the experiments leading to matrix $\Lambda$ with the smallest condition number) can be considered as optimal among the set of subsets considered.

More details about this approach, as well as some applications can be found in the works of Marchuk (1994) and Marchuk and Agoshkov (1993).

## 6. Optimization of Emissions from Enterprises by Using Some Economical Criteria

Here we consider the problem of determining the quantity of enterprise emissions, such that the concentrations of any pollutant of a given region will be less than or equal to the pollution level which ensures that the ecosystem will not be destroyed. Another problem which is very close to this one is the problem of optimization of the location of new enterprises so that the pollution of a given region will not exceed prescribed critical levels which are damaging for plants, animals, and humans. We are trying to determine emission levels which do not damage the ecosystem, but are not too low, in order to achieve the maximum economical efficiency for the given restrictions. It is important to consider such kinds of optimization problems especially for Central and East European countries, which must quickly achieve sustainable development of their industry and agriculture.

Let us consider the region $\Omega$ and its boundary $\partial\Omega$. Let $n$ enterprises (or sources of pollutants) $A_i$ be located in the points $r_i$, $i = 1, 2, ..., n$. Assume that all sources can be considered as point sources $E_i$ ($i = 1, 2, ..., n$). Introduce $m$ ecological domains $\omega_k$, $k = 1, 2, ..., m$. Suppose that the limits of the pollution levels for the integrated interval of time $[0, T]$ are defined for every ecological domain.

In this case the model of air pollution transport for a fixed $s$th pollutant (1) can be presented in the following form

$$\frac{dc(r,t)}{dt} - A(r,t)c(r,t) = \sum_{i=1}^{n} E_i \delta(r - r_i). \tag{45}$$

Consider the functional induced by the domain $\omega_k$.

$$Y_{sk} = \int_0^T dt \int_{\omega_k} p_c c_s d\omega, \quad k = 1, ..., m. \tag{46}$$

Thus, functional (46) describes the admissible pollution levels in the ecodomain $\omega_k$. The problem is to find the set of planned emissions $E_i$ ensuring

$$Y_{sk} \leqslant (C_s)_k, \quad k = 1, 2, ..., m, \tag{47}$$

for minimal economical losses for technical reconstruction of the enterprises, where $(C_s)_k$ is the critical level of the considered $s$th pollutant for the domain $\omega_k$.

It is clear that in this case we have to consider the minimizing functional

$$I = \int_{i=1}^{n} \xi_i(\hat{E}_i - E_i), \tag{48}$$

together with constraints (47), where $\hat{E}_i$ and $E_i$ are the initial and planning power of the emissions, $\xi_i$ describes the economical investments in technologies needed to achieve the same level of production with the reduced emissions (calculated per unit of emission power). Then the functional (48) describes the losses needed to improve the technologies of all $A_i$ enterprises in order to switch from emission powers $\hat{E}_i$ to the planned emission powers $E_i$. This means that we can formulate the following **minimization problem**: find the power of the sources $E_i$ in model (45) such that

$$I = \int_{i=1}^{n} \xi_i(\hat{E}_i - E_i) = \min, \tag{49}$$

$$Y_{sk} \leqslant (C_s)_k, \quad k = 1, 2, ..., m. \tag{50}$$

The problem (45), (49), (50) can be considered as a linear programming problem (Gill *et al.*, 1982). There are two approaches to solving this problem: the first one uses the original equations, the second one is based on the theory of the adjoint equations. The approach which is based on the theory of the adjoint equations can be derived by using a technique which is similar to that presented in Section on Keeping the Concentrations in a Given Sensitive Area under a Prescribed Level.

**ACKNOWLEDGMENTS**

This study was partially supported by grants from NATO (OUTR. CGR.960312), the European Union (WEPTEL Esprit 22727 and EUROAIR Esprit 24618), and the Bulgarian Ministry of Education, Science and Technology (I-811/98). A grant from the Danish Natural Sciences Research Council gave us access to all Danish supercomputers.

## References

Agoshkov, V. 1994. "Control theory approaches in: data assimilation processes, inverse problems and hydrodynamics." In *HERMIS'94*, edited by E. A. Lipitakis Vol. 1: pp. 21–39. Helenic Mathematical Society.

Austin, J. 1992. "Toward the four dimensional assimilation of stratospheric chemical constituents." *J. Geophys. Res.* **97(D2)**: 2569–2588.

Bahvalov, N. S. 1964. "On the optimal estimations of convergence of the quadrature processes and integration methods." In *Numerical methods for solving differential and integral equations,* pp. 5–63. Nauka, Moscow, Russia.

Bertsekas, D. P. 1982. In *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, New York, USA.

Bitzadze, A. V. 1982. In *Equations of the Mathematical Physics*. Nauka, Moscow, Russia.

Courtier, P., and O. Talagrand. 1987. "Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results." *Q. J. R. Meteorol. Soc.* **113**: 1329–1347.

Curtiss, J. H. 1954. "Monte Carlo methods for the iteration of linear operators." *J. Math Phys.* **32(4)**: 209–232.

Curtiss, J. H. 1956. "A theoretical comparison of the efficiencies of two classical methods and a Monte Carlo method for computing one component of the solution of a set of linear algebraic equations." In *Proceedings of Symposium on Monte Carlo Methods,* pp. 191–233. John Wiley and Sons.

Derber, J. C. 1985. *The variational four-dimensional assimilation of analyses using filtered models as constraints*. PhD dissertation. Madison, Wisconsin, USA: University of Wisconsin.

Dimov, I., U. Jaekel, and H. Vereecken. 1996. "A numerical approach for determination of sources in transport equations." *J. Comput. Math. Appl.* **32(5)**: 31–42.

Dimov, I., U. Jaekel, H. Vereecken, and D. Wendt. 1997. "A numerical approach for determination of sources in reactive transport equations, Second International Conference on Monte Carlo and Quasi-Monte Carlo methods in scientific computing, University of Salzburg, 8–12 July, 1996." In *Proceedings of MC & QMC 96, Springer Notes in Statistics*, edited by H. Niederreiter, P. Hellekalek, G. Larcher, and P. Zinterhof pp. 189–204. Springer-Verlag, Berlin, Germany.

Dimov, I., and O. Tonev. 1989. "Monte Carlo methods with overconvergent probable error." In *Numerical Methods and Applications, Proceedings of International Conference on Numerical Methods and Applications, Sofia* pp. 116–120. Publishing House of Bulgarian Academy of Science, Sofia, Bulgaria.

Dimov, I., and O. Tonev. 1993a. "Random walk on distant mesh points Monte Carlo methods." *J. Stat. Phys.* **70(5/6)**: 1333–1342.

Dimov, I., and O. Tonev. 1993b. "Monte Carlo algorithms: performance analysis for some computer architectures." *J. Comput. Appl. Math.* **48**: 253–277.

Elbern, H., H. Schmidt, and A. Ebel. 1997. *Variational data assimilation for tropospheric chemistry modeling*. Köln, Germany: Institute of Geophysics and Meteorology, University of Cologne. January 3 No 97.0030.

Ermakov, S. M., and G. A. Mikhailov. 1982. In *Statistical Modeling*. Nauka, Moscow, Russia.

Ermakov, S. M., V. V. Nekrutkin, and A. S. Sipin. 1984. In *Random Processes for Solving Classical Equations of Mathematical Physics.* Nauka, Moscow, Russia.

Gery, M. W., G. Z. Whitten, J. P. Killus, and M. C. Dodge. 1989. "A photochemical kinetics mechanism for urban and regional modeling." *J. Geophys. Res.* **94**: 12925–12956.

Gill, P. E., W. Murray, and M. H. Wright. 1982. In *Practical Optimization*. Academic press, London, UK.

Hammersley, J. M., and D. C. Handscomb. 1964. In *Monte Carlo Methods*. John Wiley & Sons, Inc., New York, USA.

Hollingsworth, A., A. C. Lorenc, M. S. Tracton, K. Arpe, G. Cats, S. Uppala, and P. Kallberg. 1985. "The response of numerical weather prediction systems to FGGE level IIb data. Part I: Analysis." *Q.J.R. Meteorol. Soc.* **111**: 1–66.

Kantorovich, L. V., and G. P. Akilov. 1977. In *Functional analysis*. Nauka, Moscow, Russia.

Le Dimet, F. X., and O. Talagrand. 1986. "Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects." *Tellus* **38A**: 97–110.

Lewis, J. M., and J. C. Derber. 1985. "The use of adjoint equations to solve a variational adjustment problem with advective constraints." *Tellus* **37A**: 309–322.

Lions, J. L. 1971. In *Optimal Control of Systems Governed by Partial Differential Equations*. Springer-Verlag, Berlin, Germany.

Lorenc, A. C. 1981. "A global three-dimensional multivariate statistical interpolation scheme." *Monthly Weather Rev.* **109**: 701–721.

Marchuk, G. I. 1974. In *Numerical Solution of the Problem of Dynamics of the Atmosphere and the Ocean*. Gidrometeoizdat, Leningrad, Russia.

Marchuk, G. I. 1980. In *Methods of Computational Mathematics*. Nauka, Moscow, Russia.

Marchuk, G. I. 1985. In *Mathematical Modeling for the Problem of the Environment, Studies in Mathematics and Applications, No. 16*. North-Holland, Amsterdam, The Netherlands.

Marchuk, G. I. 1994. "Applications of adjoint equations to problems of global change and environmental protection." In *HERMIS'94*, edited by E. A. Lipitakis Vol. 1: pp. 1–20. Helenic Mathematical Society.

Marchuk, G. I., and V. I. Agoshkov. 1993. "On solvability and numerical solution of data assimilation problems." *Russ. J. Numer. Math. Math. Model.* **8**: 1–16.

McPherson, R. D., K. H. Bergman, R. E. Kistler, G. E. Rasch, and D. S. Gordon. 1979. "The NMC operational global data assimilation scheme." *Monthly Weather Rev.* **107**: 1445–1461.

McRae, G. J., W. R. Goodin, and J. H. Seinfeld. 1984. "Numerical solution of the atmospheric diffusion equations for chemically reacting flows." *J. Comput. Phys.* **45**: 1–42.

Megson, G., V. Aleksandrov, and I. Dimov. 1994. "Systolic matrix inversion using a Monte Carlo method." *J. Parallel Alg. Appl.* **4(1).**

Mikhailov, G. A. 1970. "A new Monte Carlo algorithm for estimating the maximum eigenvalue of an integral operator." *Dokl. Acad. Nauk SSSR* **191(5)**: 993–996.

Mikhailov, G. A. 1987. In *Optimization of the ''Weight'' Monte Carlo Methods*. Nauka, Moscow, Russia.

Penenko, V. V., and N. N. Obraztsov. 1976. "A variational initialization model for the fields of meteorological elements." *Sov. Meteorol. Hydrol.* **11**: 1–11.

Skelboe, S., and Z. Zlatev. 1997. "Exploiting the natural partitioning in the numerical solution of ODE systems arising in atmospheric chemistry." In *Numerical Analysis and Its Applications*, edited by L. Vulkov, J. Wasniewski, and P. Yalamov pp. 458–465. Springer, Berlin, Germany.

Sobol, I. M. 1973. In *Monte Carlo numerical methods*. Nauka, Moscow, Russia.

Talagrand, O., and P. Courtier. 1987. "Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory." *Q.J.R. Meteorol. Soc.* **113**: 1311–1328.

Tikhonov, A. N., and A. A. Samarski. 1977. In *Equations of the Mathematical Physics*. Nauka, Moscow, Russia.

Zlatev, Z. 1995. In *Computer Treatment of Large Air Pollution Models*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Zlatev, Z., I. Dimov, and K. Georgiev. 1996. "Three-dimensional version of the Danish Eulerian Model." *Z. Angewandte Math. Mech.* **76(S4)**: 473–476.